OPEN ACCESS

KING'S College LONDON

King's Research Portal

DOI:
10.1109/TMI.2012.2227337

Document Version
Peer reviewed version

Link to publication record in King's Research Portal

Citation for published version (APA):
Varnavas, A., Carrell, T., & Penney, G. (2013). Increasing the Automation of a 2D-3D Registration System. *Ieee Transactions on Medical Imaging*, *32*(2), 387-399. https://doi.org/10.1109/TMI.2012.2227337

Citing this paper
Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

General rights
Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

•Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
•You may not further distribute the material or use it for any profit-making activity or commercial gain
•You may freely distribute the URL identifying the publication in the Research Portal

Take down policy
If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 05. May. 2023

Increasing the Automation of a 2D–3D Registration System

Andreas Varnavas, Tom Carrell, Graeme Penney

***Abstract*—Routine clinical use of 2D-3D registration algorithms for Image Guided Surgery remains limited. A key aspect for routine clinical use of this technology is its degree of automation, i.e. the amount of necessary knowledgeable interaction between the clinicians and the registration system. Current image-based registration approaches usually require knowledgeable manual interaction during two stages: for initial pose estimation and for verification of produced results.**

We propose four novel techniques, particularly suited to vertebra-based registration systems, which can significantly automate both of the above stages. Two of these techniques are based upon the intraoperative "insertion" of a virtual fiducial marker into the preoperative data. The remaining two techniques use the final registration similarity value between multiple CT vertebrae and a single fluoroscopy vertebra. The proposed methods were evaluated with data from 31 operations (31 CT scans, 419 fluoroscopy images). Results show these methods can remove the need for manual vertebra identification during initial pose estimation, and were also very effective for result verification, producing a combined True Positive Rate of 100% and False Positive Rate equal to zero. This large decrease in required knowledgeable interaction is an important contribution aiming to enable more widespread use of 2D-3D registration technology.

***Index Terms*—2D-3D registration, registration verification, image guided surgery**

I. Introduction

Over the past decade a large number of image-based 2D-3D registration techniques have been proposed to facilitate Image Guided Surgery (IGS) in a wide area of procedures, such as orthopaedic surgery, interventional neurosurgery and vascular surgery [1]. Such techniques aim to provide a link between preoperative 3D data (e.g. Computed Tomography (CT)) and intraoperative 2D images (e.g. fluoroscopy), to provide surgeons with additional 3D information. However, despite the large number of suggested techniques, routine use of this technology in clinical practice is limited. In a recent review of 2D-3D registration techniques [1], out of over 200 referenced studies, only two refer to commercial systems incorporating such technology[2], [3] and these are both for radiation therapy applications. Overviews of computer assisted spine surgery state the field is in its infancy and that relevant techniques are initially likely to be time-consuming and have an associated steep learning curve [4], and that interaction between the surgeon and the navigation system can be a source of inaccuracy [5]. Systems which use mechanical tracking to carry out 2D-3D machine-based registration are used in cardiac [6] and neurosurgical applications [7]. In both cases manual interaction is required, either for initial positioning or verification, and significant patient movement will necessitate re-registration.

A. Varnavas, T. Carrell and G. Penney are with the Department of Biomedical Engineering, Division of Imaging Sciences and Biomedical Engineering, King's College London, King's Health Partners, St. Thomas' Hospital, London, UK.

T. Carrell is also with the Department of Vascular Surgery, Guy's and St Thomas' NHS Foundation Trust, King's Health Partners, London, UK.

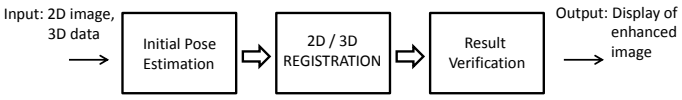

Fig. 1. The stages of a 2D/3D registration system used in clinical practice.

We believe that *automation* is a key aspect, essential for the wider establishment of 2D-3D registration systems in clinical practice. By the term *automation*, we mean all those functionalities that enable clinical use of the 2D-3D system, without it requiring *knowledgeable* input from the clinicians during the operation. The stages of a 2D-3D registration system are presented in Figure 1. At present, knowledgeable input is usually required in two main stages of the system's use: the initial pose estimation of the 3D data and the verification of the produced output. We propose here four techniques that can minimise the required interaction with the system during these stages. These techniques are primarily related to vertebra-based 2D-3D registration. However, the proposed techniques can find application under a more general framework.

Vertebra-based 2D-3D registration techniques have been widely proposed [8], [9], [10], [11], [12], [13], [14], [15], [16], [17], [18] as a means of enabling image guided surgery in the area of the spine. This is natural as vertebrae are clearly observed features on fluoroscopy for most abdominal or thoracic interventions. However, image-based registration algorithms require an initial pose estimation for the 3D data. This pose has to be estimated during the system's initialisation stage and automation of this process is expected to have multiple benefits: time saving, reduced training and reduced user errors. Most of the references above do not tackle how to obtain an initial estimate of the 3D pose. They simply evaluate algorithm performance starting from a variety of different poses [15], [8], [9], [10], [11], [12], [13], [14]. In clinical practice the required initial pose of the 3D data is usually found through a process of manually selecting anatomical landmarks in the preoperative and intraoperative images [16], [17], [18]. This severely complicates the system's use, and for vertebra-based systems, identifying specific vertebra in fluoroscopy can be a very difficult task even for highly experienced clinicians [19].

Outside the vertebra-based registration literature, a technique for robust initial pose estimation is presented in [20]. The authors propose a spectrum based similarity measure for identifying the rotation of the 3D data, combined with phase correlation for the estimation of the in-plane translation. In experiments on clinical patient data (registering a 3D rotational x-ray of a skull to a 2D x-ray) their method achieved a success rate of 68.6% (compared to 28.8% when the robust initialization is not used) for starting positions up to 21mm from registration. Starting positions for our application would most likely have errors greater than 21mm (an incorrectly chosen vertebra results in approximately a 30mm error). In addition we require a method to work during interventional procedures where the preoperative image could be a CT scan of very different field of view to the fluoroscopy image, and the fluoroscopy images may be low-dose (i.e. noisy) images which contain many additional features such as bowel gas and interventional instruments. All these factors are likely to reduce the similarity between the 3D and 2D images, and so reduce the performance of an initial pose estimation method. For these reasons we believe the method proposed by van der Bom *et al.* [20] would not have a sufficient capture range for our application.

The literature with respect to verifying registration results in clinical practice is also limited. The typical experimental approach is to evaluate registration results using fiducial markers attached on phantoms or cadavers [21]. However, the use of accurate fiducial markers (e.g. bone implanted) is rarely acceptable for widespread clinical applications because additional surgery is required for their placement. Visual assessment through a process of superimposing, on the fluoroscopy image, a Digitally Reconstructed Radiograph (DRR[1]) corresponding to the pose found, is an alternative way of evaluating the result of the registration process and is mentioned in [15], [17], [18]. This, however, requires a certain amount of experience for accurate assessment and can significantly delay the operation when the registration results need to be updated frequently.

We present here four novel techniques that can be used to increase the automation of a vertebra-based 2D-3D registration system. Two of them address the initial pose estimation of the 3D data (through vertebra identification) and the other two tackle the problem of verifying the registration results. Two of the proposed techniques are based upon the generation of a reference point which is physically present only in the intraoperative 2D data. The 3D preoperative coordinates of the point, named as Virtual Fiducial Marker (VFM), are computed using initial registrations of the 2D and 3D data. We show how a VFM can be used to: 1) identify a vertebra of interest (Section II-B1) and 2) generate an index which is correlated with the success of the registration process (Section II-C1). The remaining two proposed techniques use an inter-vertebrae registration scheme, based upon the idea of registering a 2D area of interest to a number of different 3D vertebrae. We show how the produced similarity values can be used to: 1) find the correct vertebra corresponding to the area of interest (Section II-B2) and 2) detect misregistrations (Section II-C2).

Once a registration has been achieved and verified, some image-guided surgery systems track the fluoroscopy system either by mechanical [7] or optical [15] tracking, or via fiducial markers [22]. This enables a registration to be maintained when the fluoroscopy system is repositioned. However, patient movement will result in registration errors, which can require re-registration. Other systems re-register whenever the fluoroscopy set is moved [23] and any patient movement is accounted for by the new registration. In image guided orthopaedic procedures it is common to rigidly attach a tracked marker to the bone of interest, which allows the registration to be maintained to the tracked bone despite patient movement. Our proposed techniques to increase automation are most useful in systems where repeated registrations are required (either due to fluoroscopy system or patient movement) or when vertebra identification is particularly difficult (e.g. when using a fluoroscopy set with a small field of view, or patients with abnormal anatomy).

This paper is organised as follows. In Section II we present the proposed methods for automating initial 3D pose estimation and verification of the registration result. These methods are evaluated and discrimination thresholds calculated in Section III, using data from 8 complex endovascular aortic repairs (74 fluoroscopy images and 8 CT scans). In Section IV we present a retrospective case study: investigating the impact of the proposed methods on the performance of an Image Guided System which was used during 23 complex aortic aneurysm repairs (*fenestrated or branched endovascular grafts*)[18]. The discussion of the results and the conclusions follow in Sections V and VI respectively.

II. Methods

A. Virtual Fiducial Markers (VFMs)

When Fiducial Markers are used to register preoperative data to the intraoperative scene they must be placed on or into the patient before acquiring the 3D data and they have to remain at the same position until the operation takes place. This requirement makes the process of placing fiducial markers particularly demanding; generally additional surgery has to be undertaken. As such, fiducial markers are usually used only when they are necessary for the registration process, and not just for complementary tasks, such as initial pose estimation or result verification.

We propose here the concept of a Virtual Fiducial Marker (VFM), a technique that enables us to "insert" a reference point in the preoperative 3D data at the time of the operation. This can be achieved when two successful 2D-3D registrations, under an appropriate relative fluoroscopy gantry movement, have been produced. The user can then select a characteristic point (VFM) appearing in both 2D images and compute its 3D coordinates in the coordinate system of the preoperative data using *triangulation*. Let us also note here that the points selected as VFMs should not move relative to the patient during the operation.

The idea of the VFM is based on the observation that there is a big variance in the difficulty of initialising and verifying

[1]A DRR is a synthetic x-ray image produced by projecting in-silico rays through the CT image and integrating the Hounsfield values

an image-based registration during a medical procedure. This is due to differences in the quality of the images, in the presence or absence of distinctive features or in the difficulty of identification of pictured anatomy. We propose acquisition of two high quality, wide field of view images, separated by a 20° lateral rotation, at the start of the operation. Wide field of view images, which contain distinctive landmark features, such as the pelvis and lowest rib, allow the registration process to be easily initialised, and the subsequent registrations to be easily verified through visual inspection. The produced registrations are used to construct the VFM, after which the procedure can continue following standard imaging protocols and the VFM can be used to help the image guidance system when more challenging images are acquired. The information contributed by a VFM can be used in the process of initial pose estimation and registration result verification.

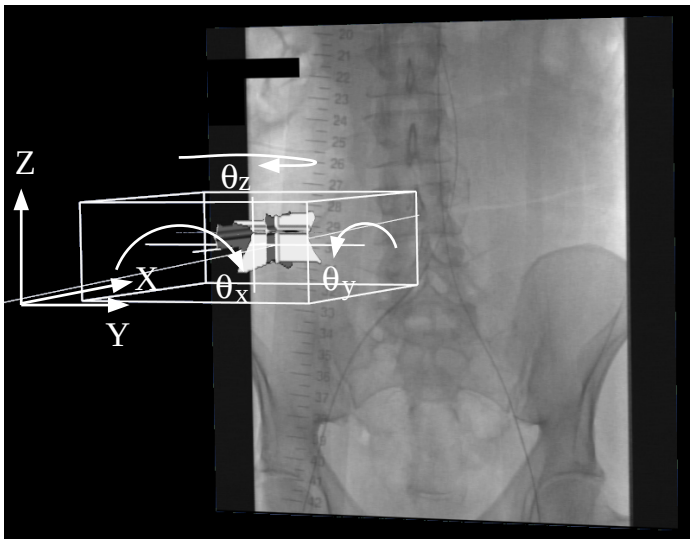

Fig. 2. CT position and orientation is defined by six rigid body parameters, three translations X, Y, Z, and three rotations $\theta_X, \theta_Y, \theta_Z$. These can be split into parameters which define movements parallel to the plane of the fluoroscopy image (in-plane parameters θ_X, Y, Z) and parameters which define movements a component of which is normal to the fluoroscopy plane (out-of-plane parameters θ_Y, θ_Z, X).

B. Techniques for initial pose estimation

A typical, image-based, 2D-3D registration system requires an initial estimate of the pose of the 3D data. The accuracy needed varies depending on the registration algorithm and quality of image data, however the closer the initial estimate is to the true pose of the 3D data, the higher is the probability of producing a correct registration. We present in Figure 2 the six parameters that can be used to describe the pose of the 3D data. Some of these parameters can be easily estimated to within the capture range of a registration algorithm using knowledge of patient orientation and information easily accessible from most fluoroscopy sets. For example, when patients are imaged in a supine position for both CT and fluoroscopy, then the rotational parameters can be estimated using the standard information on gantry orientation provided by nearly all fluoroscopy sets. The accuracy of these estimates will depend on how reproducibly the patient lies in a supine position (3-4° on average, max 8° in our experience [18]), and on the accuracy of fluoroscopy gantry tracking. In addition although the out-of-plane translation parameter is difficult to accurately calculate, 2D-3D algorithms often have a large enough capture range for this parameter that a constant value can be used as a starting estimate (800mm for all our experiments). The most difficult parameters to estimate, which are not available on any (except for very high-end fluoroscopy sets), are the in-plane translation parameters (Y and Z in Figure 2), which can often be altered by either moving the fluoroscopy set, or the patient bed. These parameters are usually obtained by manually identifying anatomical landmarks on the fluoroscopy image.

When the registration system is being used in the area of the spine, the process described above usually takes place through vertebra identification. Identification of specific vertebrae can be very challenging if anatomy such as the pelvis or lowest rib are not clearly visible. This can often occur when high fluoroscopy magnification settings are used. In this section we present two methods that can be used to help the user identify the vertebra of interest and thus produce an initial estimate of the in-plane translation of the 3D data.

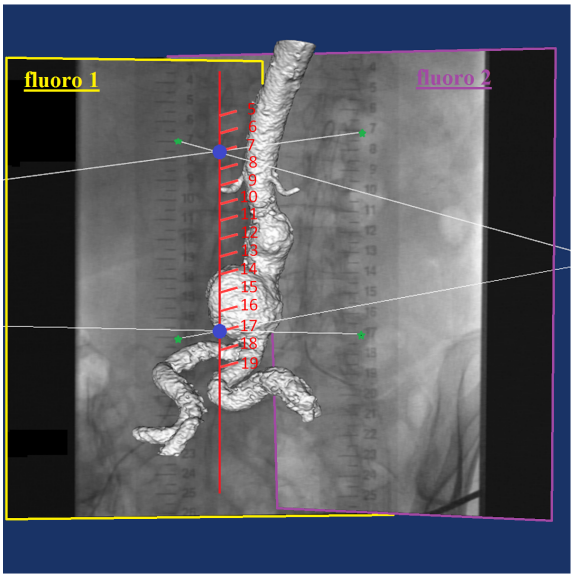

Fig. 3. The 3D surface of the aorta (from the preoperative CT scan) and two intraoperative fluoroscopy images are brought into spatial agreement through two separate 2D-3D registrations. The user selects corresponding ruler points from each 2D fluoroscopy image (e.g. point 7 denoted with a green star). The corresponding 3D coordinate of these points is calculated through triangulation (denoted as a blue circle). This process is repeated a second time (e.g. point 17) and the corresponding 3D point is calculated. The 3D position of the entire ruler can then be determined from these two 3D points by interpolation and extrapolation (denoted by red line).

1) Vertebra identification using a Virtual Fiducial Marker: We show here how a single VFM can be used to help in the process of vertebra identification. Let us note first that we construct a number of VFMs at the counting points of a

radiopaque ruler, which is fixed to the operating table during the operation. Two points of the ruler have to be visible in both images used for the reconstruction of their 3D position in the CT coordinate system (see Figure 3). The 3D coordinates of the rest of the points can then easily be extrapolated or interpolated. A VFM can then be used at any point during the operation by selecting a point on the ruler in the fluoroscopy image and indicating the number associated with it.

Let us denote with x, y, z the 3D coordinates of the CT and with $\tilde{x}, \tilde{y}$ the 2D coordinates of the fluoroscopy image. Let us also assume that the origin of the 2D coordinate system is at the principal point of the fluoroscopy plane. Then, the projection of a 3D point on the fluoroscopy image can be described by:

$$\boldsymbol{K}[\boldsymbol{R}|\boldsymbol{t}]\boldsymbol{x} = \lambda\tilde{\boldsymbol{x}} \tag{1}$$

,where $\boldsymbol{x} = \begin{pmatrix} x & y & z & 1 \end{pmatrix}^T$ and $\tilde{\boldsymbol{x}} = \begin{pmatrix} 1 & \tilde{x} & \tilde{y} \end{pmatrix}^T$ are the homogenous representations of the 3D and 2D points respectively, $\boldsymbol{R}$ is the 3×3 rotation matrix, $\boldsymbol{t} = \begin{pmatrix} t_x & t_y & t_z \end{pmatrix}^T$ is the 3×1 vector denoting the translation component, $\boldsymbol{K} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & f \end{pmatrix}$ is the 3×3 calibration matrix of the fluoroscopy set, f is the focal length of the fluoroscopy set and λ is a scalar related to the normalisation of the homogenous vector $\tilde{\boldsymbol{x}}$.

Let us now assume that $\boldsymbol{x}$ and $\tilde{\boldsymbol{x}}$ are the known 3D and 2D coordinates of a VFM determined by triangulation, and that the system has an initial estimation for the rotation and the out-of-plane translation of the CT data, i.e. matrix $\boldsymbol{R}$ and t_x are known. One can then use this information to acquire an estimate for the in-plane translations t_y and t_z by rewriting Eq.1 as:

$$\begin{pmatrix} 1 & 0 & 0 \\ \tilde{x} & -1 & 0 \\ \tilde{y} & 0 & -1 \end{pmatrix} \begin{pmatrix} \lambda \\ c_y \\ c_z \end{pmatrix} = \boldsymbol{K}\boldsymbol{R} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} c_x \\ 0 \\ 0 \end{pmatrix} \tag{2}$$

, where:

$$\begin{pmatrix} c_x & c_y & c_z \end{pmatrix}^T = \boldsymbol{K}\boldsymbol{t} \tag{3}$$

We note here that because of the form of matrix $\boldsymbol{K}$, c_x is known, whereas c_y and c_z are unknowns. Once Eq. 2 is solved with respect to $\begin{pmatrix} \lambda & c_y & c_z \end{pmatrix}^T$, the in-plane translations t_y and t_z can be computed using Eq. 3.

Therefore, by using a single VFM position and estimates for the rotation and out-of-plane translation parameters we can obtain a registration starting position. This is used to overlay the chosen CT vertebra onto the fluoroscopy image. The accuracy of the overlay will depend on the accuracy of the estimated parameters. To increase accuracy the user can fine tune the estimate by manually translating the overlaid CT vertebra (by altering the in-plane translation parameters) to overlay the closest fluoroscopy vertebra. If the estimated parameters have large errors, it is possible that the closest fluoroscopy vertebra will be incorrect. This possibility is investigated in our evaluation of methods, section III-A1.

2) Inter-vertebrae optimisation based on a similarity value: The difficulty of estimating the in-plane translation through vertebra identification lies in the fact that vertebrae look very similar one to the other. However, vertebrae in general have a distinctive shape and one can easily identify their projection on a fluoroscopy image. Based on these two observations, we propose here a semi-automated method for the initial estimation of the in-plane translation through automated vertebra identification.

The method requires the user to draw a bounding box around the projection of a vertebra on the fluoroscopy image. Note that the corresponding vertebra from the CT scan does not need to be identified. Instead a 2D-3D registration is carried out between a number (e.g. T10, T11, T12, L1, L2, L3, L4 and L5) of likely candidate CT vertebrae. The output from the registration algorithm is the pose of each vertebra, and a final value for the similarity measure between the CT vertebra and fluoroscopy vertebra within the bounding box. Our algorithm uses a modified version of the Gradient Difference similarity measure [23] as described by Eq. 5, which is very sensitive to both changes in vertebra pose and shape. Consequently the vertebra which results in the highest similarity value can be used to identify the correct corresponding vertebra. Therefore, our 2D-3D registration optimises seven parameters in total: the six rigid body pose parameters and the vertebra name.

C. Techniques for automated verification of registration result

As mentioned earlier the ability of a 2D-3D registration system to provide feedback to the user concerning how well the registration process has worked is critical, particularly when it is used as part of a clinical image guided system. We propose in this section two methods that can be used to provide indices of how well the registration process has worked and so aid detection of misregistrations.

1) Verification of registration result using a Virtual Fiducial Marker: We show here how a single VFM can be used to generate an index which is correlated with the accuracy of the registration process. Let us note here that a number of VFMs are initially "inserted" into the 3D CT data using 2D positions picked on a radiopaque ruler visible in the fluoroscopy images, as described in Section II-B1. Once the registration process has been completed, the user can select a point on the ruler, indicating the 2D fluoroscopy image coordinates of the VFM $\widetilde{\boldsymbol{x}} = \begin{pmatrix} \tilde{x} & \tilde{y} \end{pmatrix}^T$. We propose that the reprojection distance between $\widetilde{\boldsymbol{x}}$ and its 3D coordinates $\boldsymbol{x} = \begin{pmatrix} x & y & z \end{pmatrix}^T$ transformed using the parameters produced by the registration process (i.e. to position $\boldsymbol{x}_{rt} = \boldsymbol{R}\boldsymbol{x} + \boldsymbol{t}$, $\boldsymbol{R}$ and $\boldsymbol{t}$ are the rotation and translation matrices respectively) can be used as a verification index. Assuming a 3D fluoroscopy coordinate system, with the origin at the x-ray source, the suggested reprojection distance d is the minimum distance of the point $\boldsymbol{x}_{rt}$, to the ray from the xray source to position $\widetilde{\boldsymbol{x}}$ on the 2D image. This distance can be computed as:

$$\begin{aligned} d &= \|\boldsymbol{x}_{rt}\| sin(\phi) \\ \phi &= arccos\left(\frac{\tilde{\boldsymbol{x}}_f{}^T \boldsymbol{x}_{rt}}{\|\tilde{\boldsymbol{x}}_f\|\|\boldsymbol{x}_{rt}\|}\right) \end{aligned} \tag{4}$$

, where $\tilde{\boldsymbol{x}}_f = \begin{pmatrix} f & \tilde{x} & \tilde{y} \end{pmatrix}^T$ and f is the focal length.

The reprojection distance, d Eq. (4), is the minimum distance that the VFM has to be moved in 3D space (after registration) in order to agree with its projection pictured on the fluoroscopy image. Zero reprojection distance does not mean zero registration error, but if the registration errors are random, then there should be a strong positive correlation between the value of d and registration accuracy. The use of multiple VFMs for validation should further increase the strength of this correlation.

2) Verification of registration result using normalised similarity values: Image-based 2D-3D registration algorithms search for the maxima of a similarity value defined in the space of the registration parameters. The value of the similarity measure can depend on a number of factors:

1) relative pose of 2D and 3D images
2) additional features (e.g. surgical instruments) in 2D field of view.
3) 2D image quality (noise, artifacts)
4) 3D image quality (noise, artifacts)

For a given registration nearly all these factors are held constant, apart from the registration parameters which define the relative pose of the two images. Therefore, 2D-3D registration can still be achieved by maximising the similarity measure. However, the final value of the similarity measure will still be effected by all the above factors, and so using the final value of the similarity measure to define whether or not a registration has succeeded is non-trivial.

We propose designing a decision rule about the correctness of the registration process combining the similarity value at the position found (*registration value*), with the one corresponding to another maximum in the parameter space (*normalisation value*). We calculate our normalisation value by holding constant as many of the above factors as possible. We note that much larger changes occur in interventional fluoroscopy images (image quality, field-of-view, instruments) than in the CT scan. We therefore propose to calculate our normalisation value by carrying out additional registrations using the same fluoroscopy field of view, but with a different 3D CT anatomical entity.

In the case of a vertebra-based registration system, the suggested verification method can be incorporated as follows. The registration process is first performed after the user has identified a vertebra on the fluoroscopy image, and the final value of the similarity measure is saved (registration value). The registration process is then repeated a number of times, but each time an incorrect 3D vertebra is used. The maximum similarity value after matching to these incorrect vertebrae is saved (normalisation value). A two dimensional feature space is then constructed, with the first feature being the registration value and the second feature being the ratio of the registration value over the normalisation value, named *similarity ratio*. Populating this feature space with samples of successful and failed registrations, one can develop a suitable decision rule for their discrimination.

We use a modified version of the Gradient Difference [23] similarity measure to allow more meaningful comparison of similarity values between different datasets; between different fluoroscopy images of the same patient, and between different patients. Three modifications are carried out which help to reduce the effect of the following three factors on the final similarity value:

1) The size of the region of interest used for registration,
2) The distribution of gradient values in the fluoroscopy region of interest,
3) The fluoroscopy pixel size.

It should be noted that all the above factors remain constant within the optimisation process of any single registration, and so the original gradient difference measure did not need to account for them. Similarly, all three proposed modifications do not change the way that the similarity measure performs for a single registration. We first define our modified similarity measure, and then detail how it enables more meaningful comparisons across different data sets.

The modified similarity measure g between an $M \times N$ region of interest $I(x, y)$ of a fluoroscopy image and the corresponding region $J(x, y)$ of a DRR is computed according to:

$$g = \max_{s_x, s_y} \left\{ \frac{1}{MN} \sum_{i=1,j=1}^{M,N} \left(\frac{\sigma_x^2}{\sigma_x^2 + (I_x(i,j) - s_x J_x(i,j))^2} + \frac{\sigma_y^2}{\sigma_y^2 + (I_y(i,j) - s_y J_y(i,j))^2} \right) \right\} - g_0 \tag{5}$$

$$g_0 = \frac{1}{MN} \sum_{i=1,j=1}^{M,N} \left(\frac{\sigma_x^2}{\sigma_x^2 + I_x(i,j)^2} + \frac{\sigma_y^2}{\sigma_y^2 + I_y(i,j)^2} \right)$$

where I_x, I_y, J_x, J_y are the horizontal and vertical gradient fluoroscopy and DRR images computed using the Sobel operator. σ_x^2 and σ_y^2 are the variance of the horizontal and vertical gradient values in the fluoroscopy region of interest. s_x and s_y are scaling factors for the gradient DRRs and their values are optimised separately at each step of the registration optimisation, using one dimensional gradient descent search strategy [23].

This new formulation addresses the three factors stated above by:

- Division by the number of pixels MN results in our measure being normalised for the size of the region of interest. It should be noted that since this number remains constant during the registration process, the registration performance of the similarity measure remains the same.
- The subtraction of the term g_0 in Eq. 5 helps normalise the measure with respect to the gradient value distribution of the fluoroscopy image. Note that, for a single registration, the term g_0 will be constant. Therefore, subtraction of g_0 will not change the performance of the 2D-3D registration algorithm at all. The value of g_0 depends only on the region of interest defined in the fluoroscopy image, and equals the value of the original gradient similarity measure [23] when s_x and s_y equal zero. By subtracting the value g_0 the starting value (i.e. when s_x and s_y equal zero) of the similarity measure is now zero, regardless of the gradient value distribution of the fluoroscopy image. The final value of the similarity measure g then represents

the improvement in gradient difference that has been achieved by subtracting gradient information from the DRR images. This subtraction of the baseline gradient difference value g_0 helps to normalise g, and so allows more meaningful comparison between gradient difference values from different registrations.
- After registration, the fluoroscopy and DRR images are resampled to a standard pixel size, and g is recomputed. The resampling process ensures that the scale at which the gradient information is computed in Eq. 5 remains invariant with respect to the resolution of the acquired fluoroscopy image.

III. Evaluation of Methods

In this section we evaluate the performance of the proposed methods, using data from 8 complex aortic aneurysm procedures. All the CT scans used in this research were standard diagnostic images, typical voxel dimensions were $1\times1\text{mm}^2$ in-slice, with reconstructed slice thickness varying between 0.5 to 3 mm (the specific CT machine used varied depending on the referring hospital). All the fluoroscopy images were acquired on a Siemens AXIOM-Artis interventional fluoroscopy system at St Thomas' Hospital which has four magnification settings which result in a varying field of view from 16cm to 38cm.

We refer to the dataset used in this section as the "evaluation set" and it consists of:

- 8 preoperative CT scans
- 16 high quality, spot film, wide field of view images (one anterior-posterior (AP) view and one approximately 20° rotated view for each procedure) which were used to insert the ruler VFM positions into the CT scans.
- 74 fluoroscopy "test" images: these were the standard interventional images acquired during these procedures, they were predominantly low-dose (i.e. high noise) screening images.

The evaluation set is used for two purposes. The first is to test the ability of the proposed algorithms to solve the problem of initial pose estimation. This is done by evaluating how well a given vertebra is identified on the fluoroscopy image. The second purpose is to study the capability of the verification of registration result techniques to detect misregistrations. This is achieved by producing one set of correct and one set of incorrect registrations from the data. The feature space of the two classes is then constructed, visualised and discriminant rules are computed. These discriminant rules are used in section IV where we investigate the impact of our proposed methods using a much larger series of data (gathered over 23 operations).

We have used a vertebra-based 2D-3D registration system for the studies reported in this paper. This system has been described in detail in [18], however, for completeness we provide a brief summary: The system has a graphical user interface that enables an initial estimation of the pose of the CT data through the identification of a vertebra visible in the fluoroscopy image. The automated registration process uses a two stage optimisation strategy using the gradient difference similarity value [23]. In the first stage, a global optimisation process evaluates the similarity value between the fluoroscopy image and 606,375 DRRs, constructed at predefined positions around the initial estimate of the pose of the CT data. The position found is refined, at the second stage, through a simple hill climbing search strategy. Registration verification is achieved using visual inspection. The target application for this system is to aid guidance during complex aortic aneurysm procedures. Registrations are carried out to a single vertebra, and then the anatomy of clinical interest (the aorta and visceral vessels from CT) are overlaid onto the fluoroscopy image. This is valid as the aorta and vertebra remain in approximately a rigid body relationship (see [18] for further details on system accuracy).

A. Evaluation of Methods for Initial Pose Estimation

1) Vertebra identification using a Virtual Fiducial Marker: The following experiment is carried out to investigate the effect errors in starting position have upon the ability of our VFM method to identify the correct vertebra. We assume a gold standard transformation between a 3D and 2D image as depicted in Figure 4(top), where the CT scan is positioned as a standard anterior-posterior view, 800mm from the x-ray source, such that our anatomical point of interest (vertebra centre point, $\boldsymbol{x}_{GS}^{vert}$) projects to the centre of the 2D image $\widetilde{\boldsymbol{x}}_{GS}^{vert}$. We define a VFM point, $\boldsymbol{x}_{GS}^{VFM}$, to be 100mm posterior and 30mm superior to the vertebra, where a radiopaque ruler could be positioned. Using the gold standard transformation the VFM projects to 2D position $\widetilde{\boldsymbol{x}}_{GS}^{VFM}$.

We perturb the position of the 3D image, from the gold standard position, by adding errors to the rotation and out-of-plane translation parameters. A starting position is estimated from these perturbed parameters using the VFM. This is achieved as described in section II-B1, or more specifically by applying Eq. 2 which ensures that the new 3D VFM position $\boldsymbol{x}_{rt}^{VFM}$ still projects exactly onto the 2D VFM point $\widetilde{\boldsymbol{x}}_{GS}^{VFM}$. Note that the new 3D vertebra point $\boldsymbol{x}_{rt}^{vert}$ no longer projects to the centre of the image. It is this error in projection of the vertebra which we measure. Specifically we calculate the *reprojection distance* (d in Eq. 4 and depicted in Figure 4(bottom)) between the new 3D vertebra point $\boldsymbol{x}_{rt}^{vert}$ and the ray back-projected from $\widetilde{\boldsymbol{x}}_{GS}^{vert}$.

Errors were added to the three rotation and out-of-plane translation parameters by sampling in the intervals: [-10° 10°] and [-150mm 150mm], in steps of 0.5° and 5mm. To visualise the computed values of the reprojection distance from the four dimensional space, we collapsed the three dimensions describing angle error into a single value equal to the maximum of the rotational errors (which we refer to as *maximum angle error*). For each maximum angle and out-of-plane error pair of values, a range of reprojection distances will be obtained. Figure 5 shows the maximum value calculated at each pair of error values. From Figure 5 it can clearly be seen how the reprojection error increases as out-of-plane and rotation errors increase. To put these values into perspective considering our specific task: the height of an average vertebra is approximately 30mm, and so an error of greater than half this height (>15mm) could result in a misidentified vertebra.

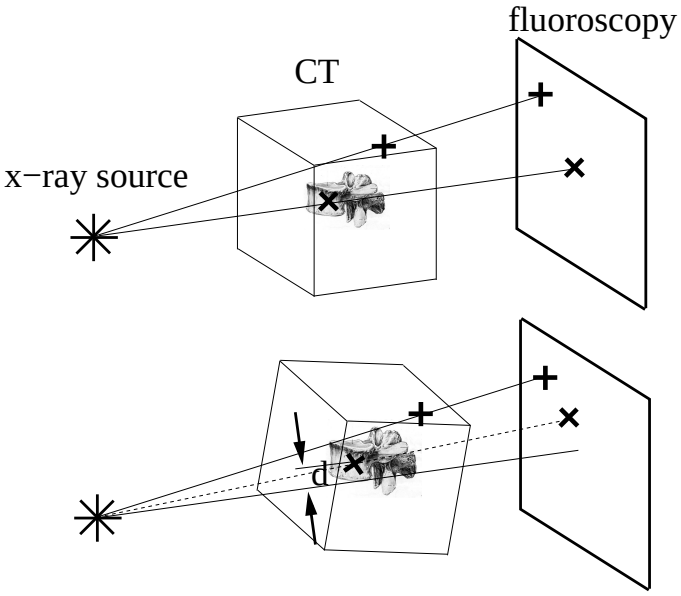

Fig. 4. Showing how error in VFM initial pose estimation was evaluated. Top shows gold standard position where vertebra point $\times$ matches to centre of fluoroscopy and VFM point $+$ position is shown. Bottom shows position after adding errors to the gold standard parameters and applying our VFM method to constrain the in-plane translation parameters, see section II-B1 for details. Note how VFM point, $+$, matches exactly, but due to rotational and out-of-plane translational errors, the vertebra no longer matches to the centre of the fluoroscopy image. Error calculated is reprojection distance d between perturbed CT vertebra position and back-projection line from gold standard 2D vertebra position.

This boundary is represented by the contour line at 15mm in Figure 5.

To determine whether sufficiently accurate starting positions could be calculated during an operation the following experiment was carried out using our 8 patient evaluation datasets. For each dataset a VFM was inserted into the CT scan using registrations to the specifically acquired spot-film AP and 20° rotated images. The initial rotational parameter values for subsequent registrations were set equal to the values from the first AP registration plus any rotational motion of the fluoroscopy gantry. The initial out-of-plane translation was set to be 800mm. A VFM ruler position was picked in each of the evaluation fluoroscopy images and used to determine the in-plane starting position values. This starting position was examined by overlaying the vertebra from CT onto the fluoroscopy image, and by visual inspection the correct vertebra was always the closest fluoroscopy vertebra to the overlay position.

After running the registration algorithm the differences between starting and final position rotational and out-of-plane translational parameters were calculated. A scatter plot of maximum rotational difference against out-of-plane translational difference is shown as the 74 dots in Figure 5. One can see that all these dots are within the region of correct vertebra identification, defined by the contour line at 15mm. In addition the reprojection distance between the centre of the target vertebra in fluoroscopy and its 3D location before registration (i.e. determined using the VFM) was calculated. The mean and max values over the 74 test images were 3.5mm and 13mm respectively. Both these results back up our visual inspection findings that in all cases the correct vertebra was identified.

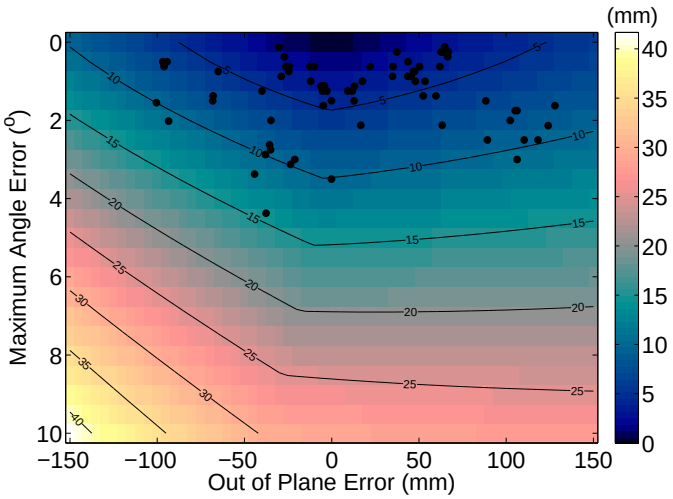

Fig. 5. Showing how the maximum reprojection distance of the centre of a chosen vertebra calculated using a VFM changes with respect to initial maximum rotation and out-of-plane translation errors. Black dots correspond to the 74 test fluoroscopy images in the evaluation set.

2) Inter-vertebrae optimisation based on similarity values: To evaluate the method proposed in Section II-B2, the registration process was performed again for 74 test images. These registrations were carried out to the known corresponding vertebra, and the final value of the similarity measure was saved, which we will refer to as *correct vertebra similarity value*. In addition the registration was repeated using a number of different CT vertebrae, up to 11 different lumbar and thoracic vertebrae were used. The best similarity value from these registrations was saved and we will refer to as *best incorrect vertebra similarity value*.

In all cases the correct vertebra similarity value was larger than the best incorrect vertebra similarity value. The median and minimum ratio of the correct vertebra similarity value divided by the best incorrect vertebra similarity value were 3.1 and 1.68 respectively. Therefore, our inter-vertebrae optimisation successfully identified the correct vertebra for all 74 test images. Moreover, the difference between the correct and best incorrect values was large, always more than 68% larger, showing the high sensitivity of our similarity measure to detect small changes in object shape.

B. Evaluation of Methods for Verification of Registration Result

To evaluate our proposed registration verification methods we produce two sets of registrations using the evaluation data set: one of correct and one of incorrect registrations. The set of correct registrations were produced by running the registration algorithm in its standard manner (described in [18]) on the 74 test images. Registration accuracy was checked using visual inspection resulting in 74 correct registrations.

Two sets of incorrect registrations were produced which will be referred to as *subset A* and *subset B*. These were produced in two different ways designed to simulate likely errors which can occur when using the algorithm.

- *subset A* represents misregistrations due to the initial position being outside the registration algorithm's capture range. These were produced by adding 20° to the initial estimates of all three angle parameters describing the pose of the CT data, and then running the registration algorithm. Such registrations were carried out to each test fluoroscopy image and visual inspection showed that 62 misregistrations occurred which form *subset A*.
- *subset B* represents misregistrations due to registering to an incorrect vertebra. These misregistrations were produced by repeating the registration process a number of times, each time using an incorrect vertebra. The result with the largest similarity value was included into *subset B*. Registrations were carried out to all 74 test fluoroscopy images resulting in 74 misregistrations in *subset B*.

1) Verification of registration using a Virtual Fiducial Marker: To evaluate the performance of the proposed VFM based verification method, we computed the VFM reprojection distance, d, as described by Eq. 4, for each of the correct and incorrect registrations. The values of d for the samples of the two classes are presented in Figure 6. We can see that there is a clear difference between the distribution for correct (mean $m = 2.57$, standard deviation $s = 2.03$) and incorrect registrations ($m' = 32.53$, $s' = 14.85$). One can exploit this fact to design a decision strategy about the correctness of the produced registration. Thus, depending on the value of the VFM reprojection distance, the result of the registration can be automatically accepted, rejected or require a manual verification by an expert.

In a system which will be routinely used in clinical practice, it is important that a verification of registration technique is used in such a way that the possibility of accepting an incorrect registration is minimised. This can be done by choosing the discriminating rule such that the False Positive Rate (FPR) is very low (preferably zero), which will usually result in a compromise having to be made with respect to the True Positive Rate (TPR). One can choose the desired trade-off between the two rates using the Receiver Operating Characteristic (ROC) curve of the proposed algorithm. We present in Figure 7 the corresponding ROC curve, which has been constructed using the samples of the two classes and varying the decision threshold based on the VFM reprojection distance (scalar value d_{VFM}).

Since we are interested in minimising the FPR, the desired point on the ROC curve at which we would like to operate, is the one where the FPR is zero and the TPR is maximum. This point is marked on the ROC curve in Figure 7 and produces a TPR of 87.84%. The corresponding discriminant threshold is shown by the horizontal line in Figure 6. Our proposed strategy for using a verification algorithm operating at such a point of the ROC curve is to accept results classified as positive by the algorithm and manually verify the ones classified as negative. Thus, since the TPR is high the majority of correct registrations will be used without the need to be verified manually by clinicians.

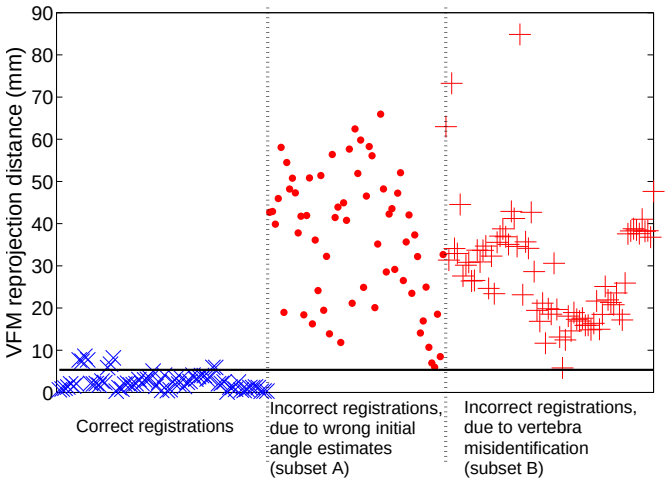

Fig. 6. Distribution of the VFM reprojection distance, used for verification of registration result. Input registrations are separated into three classes, correct registrations and misregistrations due to: algorithm finding local maxima and vertebra misidentification.

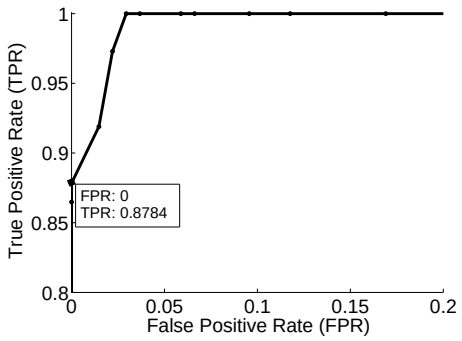

Fig. 7. The Receiver Operating Characteristic curve for the VFM based verification algorithm, constructed using the evaluation data set.

2) Verification of registration result using normalised similarity values: To evaluate the similarity value based verification method, the registration value and the similarity ratio are computed as described in Section II-C2, for each of the correct and incorrect registrations.

Due to the success of our inter-vertebrae optimisation method in clearly identifying all correct vertebrae successfully (section III-A2) we investigate here a more difficult challenge for our result verfication method: the situation when the vertebra on which the user has attempted to register on is not within the CT field of view. Such a situation could occur if a user was attempting to match to a thoracic vertebra while the CT was only of the abdomen. This is a more challenging problem because, if the correct vertebra was included, it would be such a clear winner that the misregistration would be easily identified. Therefore, when calculating the *normalisation value* for the misregistrations in *subset B*, the correct vertebra is omitted from the set of those to be considered in the

registration process.

Figure 8 shows the scatter plot of registration values and similarity ratios. We use a standard linear classifier (Fisher's linear discriminant [24]) to determine the direction of a discriminating line to separate correct registrations from misregistrations. Figure 9 shows the ROC curve produced by varying the distance of the discriminant line from the origin (scalar value $d_{NormSimVal}$) of the two dimensional feature space. As explained earlier our proposed point of interest is the one where the TPR is maximised and FPR equals zero. This point is marked in Figure 9 and produces a TPR of 94.59%. The corresponding discriminant line is drawn in Figure 8. This line is also used in section IV where we investigate the impact of this technique on a much larger set of data.

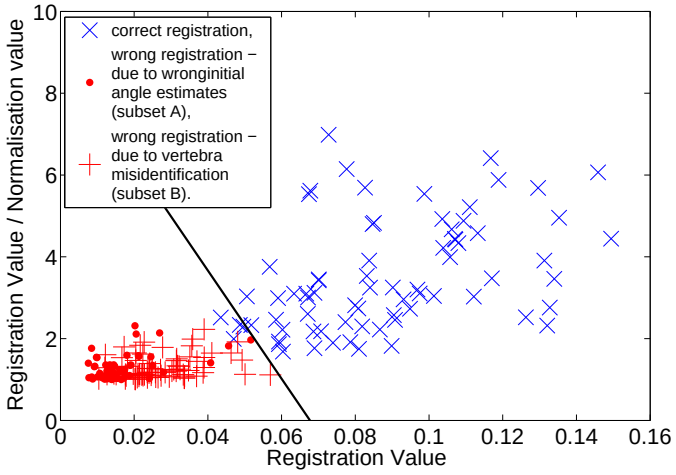

Fig. 8. Scatter plot of registration value and similarity ratio, used for verification of registration result (evaluation set).

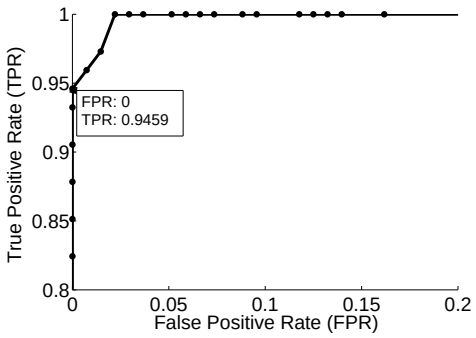

Fig. 9. The Receiver Operating Characteristic curve for the DRR similarity value based verification algorithm, constructed using the evaluation data set.

3) Verification of registration result by combining the proposed techniques: The two techniques proposed for verification of registration are based on independent criteria and as such can be combined to enhance the detection rate. This is illustrated in Figure 10 which shows each registration plotted on a scatter diagram where the two axes represent the two scalar values (d_{VFM} and $d_{NormSimVal}$) used to determine the discriminant thresholds for the VFM and normalised similarity measure techniques. For the VFM method the scalar used is simply the VFM reprojection distance. For the normalised similarity method the scalar is the distance from the origin after first projecting the data onto a line which goes through the origin and is perpendicular to the discriminant line shown in Figure 8. The horizontal and vertical lines show the values of the VFM distance value and normalised similarity threshold value determined in sections III-B1 and III-B2 to maximise TPR with FPR equalling zero.

The horizontal and vertical lines in Figure 10 split the two-dimensional space into four subspaces. Registrations falling into each of these subspaces can be described as follows:

- Upper left: registrations failed both verification tests
- Lower right: registrations passed both verification tests
- Upper right: registrations passed normalised similarity test, but failed VFM test
- Lower left: registrations failed normalised similarity test, but passed VFM test

Since the decision thresholds have been chosen such that we have zero FPR, all misregistrations are contained in the upper left subspace. Moreover, as a result of the independence of the two verification criteria, there are cases of correct registrations contained in the upper right and bottom left subspaces. One can use this observation to design a strategy that combines the two techniques to increase the TPR, while the FPR remains equal to zero. According to this strategy a registration is classified as a possible misregistration (i.e. one requiring manual verification) if it falls within the upper left subspace of Figure 10 and is classified as correct if it belongs to one of the remaining three subspaces. With respect to the data of the evaluation set, the combined verification produces a TPR of 100%. This is an improvement compared to the rates of 87.84% and 94.59% produced when each one of the proposed algorithms is used on its own, and is very close to the theoretical value of 99.34% $(1-(1-0.8784)*(1-0.9459))$ for combining two independent probabilities.

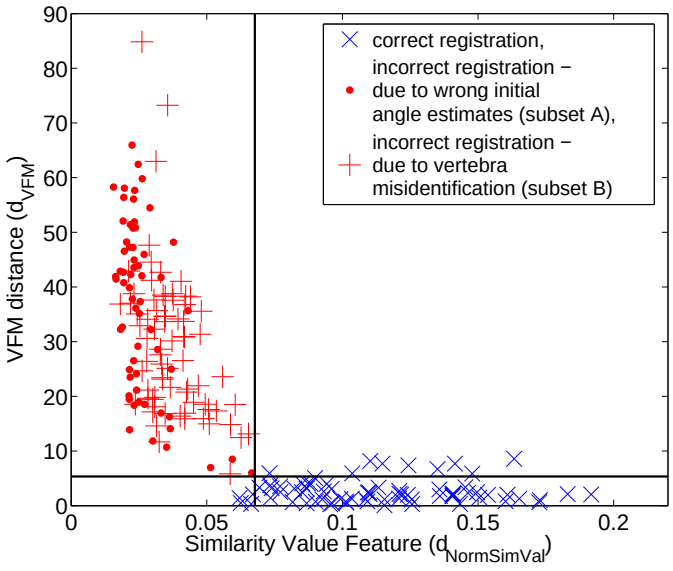

Fig. 10. Scatter plot of the features used by the two proposed verification techniques.

TABLE I
IMPACT OF THE PROPOSED TECHNIQUES ON THE RATE OF CORRECT AND WRONG (DUE TO USER ERROR AND ALGORITHMIC FAILURE) REGISTRATIONS.

Type of System	Misregistrations		Correct registration
	User Error	Algorithm failure	
Original	13 (3.7%)	7 (2.03%)	325 (94.2%)
Initial Pose Estimation (through VFM)	3 (0.87%)	7 (2.03%)	335 (97.1%)
Initial Pose Estimation (through Similarity Value)	2 (0.58%)	7 (2.03%)	336 (97.39%)

IV. IMPACT OF METHODS ON THE PERFORMANCE OF AN IMAGE GUIDED SURGERY SYSTEM

In this section we present the impact our proposed methods have on the performance of a 2D-3D registration system [18]. We use a previously acquired dataset consisting of preoperative CT data from 23 patients and 345 intraoperative fluoroscopy images which were acquired during complex aortic aneurysm procedures. The system was originally used live during these procedures (without the methods proposed here) to produce overlays of the aorta on the fluoroscopy images. The fluoroscopy images were all acquired as part of the standard clinical procedure and as such are mostly low-dose screening images. There was no overlap between the images used in this section, and those used in the evaluation dataset.

The initial pose estimation was done through vertebra identification and the results were visually verified before presented to the clinicians as described in [18]. The original performance of the 2D-3D registration system can be seen in the top line of Table I. The original system had an overall success rate of 94%, 20 failures occurred from the 345 registrations, and these have been subdivided into 13 failures due to user error (11 due to selection of incorrect vertebra, and 2 due to incorrect input of fluoroscopy gantry rotation) and 7 registration algorithm failures (finding a local maxima).

The proposed algorithms can improve our registration system in two ways.

1) Reducing misregistrations due to selection of incorrect vertebra by:
   - using the VFM
   - optimising over all vertebra
2) Providing automatic verification of registration result by:
   - using the VFM
   - using normalised value of final similarity measure

A. Impact of Methods to Reduce Misregistrations

Both the VFM and inter-vertebrae optimisation methods are able to reduce the number of misregistrations due to the user misidentifying a vertebra. To assess whether these methods would have reduced user-error misregistrations the following experiments were carried out, using all the initial correct registrations and the misregistrations due to vertebra misidentification user error.

1) Use of VFM to aid Initial Pose Estimation: Suitable images to insert a VFM into the CT data were not acquired during these operations. Therefore, we use the method described in section III-A1 to simulate the VFM effect in vertebra identification. This effect is quantified by computing the vertebra centre reprojection distance. As described in section III-A1 a VFM point is inserted into the CT scan 100mm posterior and 30mm superior to the target vertebra. The registration parameters after running the registration system (and after verification using visual inspection) were used as a gold standard final position (equivalent to Figure 4(top)). The starting position was obtained using the rotation values from the initial AP registration plus fluoroscopy gantry rotation, and assuming an 800mm out-of-plane translation.

Figure 11 shows the calculated reprojection distance for the correct registrations and the 11 registrations which failed to register due to vertebra misidentification. Results above 15mm were manually investigated by overlaying the vertebra from CT onto the fluoroscopy image using the VFM starting position parameters. All of these cases, except one (one of the previous misregistrations) would have resulted in the correct vertebra being chosen. Therefore, our VFM method should have been able to correct 10 out of 11 misregistrations caused due to vertebra misidentification. Overall system performance using the VFM is shown in the second line of Table I.

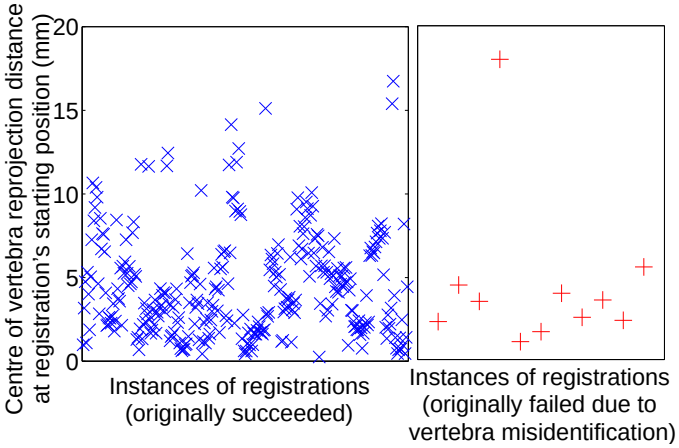

Fig. 11. The values of the reprojection distance of the centre of a chosen vertebra as resulting with the use of a VFM (Impact on IGS system).

2) Use of inter-vertebrae optimisation to aid Initial Pose Estimation: The registration process was performed again for all the previously successful registrations and the 11 registration failures due to vertebrae misidentification. The procedure presented in Section II-B2 was used, where the registration was repeated a number of times, using a different CT vertebra each time. In all cases the correct vertebra had a larger similarity value than the best value from the incorrect vertebra. Therefore, inter-vertebrae optimisation succeeded in correcting all the misregistrations due to vertebra misidentification. Again the difference between values obtained using the correct vertebra compared to the best incorrect vertebra were large. The median and minimum value of the ratio: correct vertebra value divided by best incorrect vertebra value were

2.94 and 1.28 respectively. Overall system performance using inter-vertebrae optimisation is shown in the third line of Table I.

Our proposed methods were unable to correct the two remaining failures which were due to incorrect manual input of fluoroscopy gantry rotation. This rotational information is available in the fluoroscopy system and so the obvious solution to eliminate these type of errors is an implementation which has direct access to the fluoroscopy rotational parameters.

B. Impact of Methods on Automating the Verification of Registration Result

As shown in the third column of Table I our improved system has carried out 336 successful registrations, however, 9 failures still occurred. We now investigate how successfully our automatic verification methods would be in detecting these failed registrations, and what percentage of successful registrations are automatically defined as being correct (i.e. the TPR).

We are not able to report results with the VFM based verification method, due to lack of appropriate images for VFM construction. An approach similar to the one used for the initial pose estimation would not be appropriate in this case, because the main challenge for the method is patient movement which we are unable to simulate realistically.

We therefore present here only the verification results which correspond to the use of the similarity based method presented in Section II-C2. Following the procedure described there, registration and normalisation values are produced for all images in the dataset. The similarity ratio is then computed and the scatter plot of the two features used for classification is presented in Figure 12. The discriminant straight line as computed from the data of the evaluation set in Section III-B2 is also plotted. The TPR produced by the classification algorithm is 87.8%, whereas the FPR is equal to zero.

All the misregistrations, except for one, are located well away the discriminant line in the feature space of Figure 12. The misregistration close to the discriminant line is a case when the registration algorithm found a local maximum very close to the correct registration position. As such, although the result is incorrect (determined by visual inspection), it is very close to the correct position, which was calculated by repeating the registration. The difference between the two registrations, computed as a reprojection distance of the left renal ostia (clinically relevant position) was just 2.9mm.

The percentage of registrations that would not have required manual verification (i.e. true positives over all registrations) is 85.51%. This is close to the TPR as the correct registration rate of the system is very high. Thus the majority of registrations are verified automatically which considerably improves the usability of the registration system. Let us also note here that because the discriminating criterion was chosen to detect misregistrations, the automatic verification process does not compromise the reliability of the verified results.

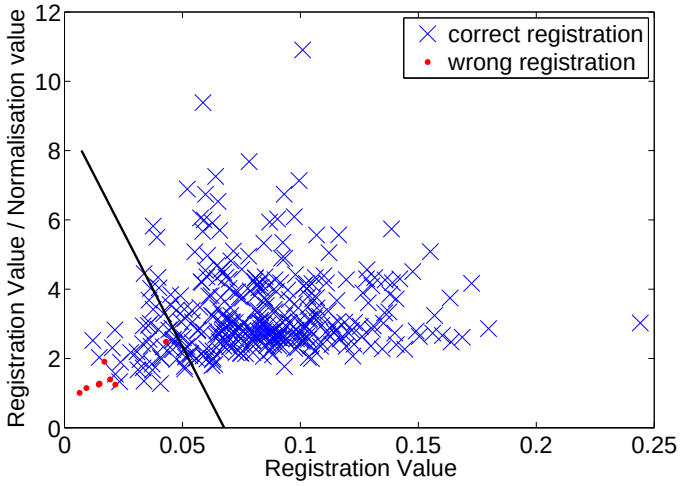

Fig. 12. Scatter plot of registration value and similarity ratio, used for verification of registration purposes (Impact on IGS system). Let us note here that there are 9 red dots in the plot (corresponding to the 9 misregistrations in the dataset), but only 7 are clearly visible due to overlapping.

V. Discussion

We presented in section IV the retrospective impact of the proposed methods on a 2D-3D registration system in terms of correct registration and verification rates. The initial pose estimation methods were able to correct almost all the misregistrations due to vertebra misidentification: the VFM method failed to correct one instance; the inter-vertebrae optimisation corrected all of these misregistrations.

For registrations which our pose estimation methods were not able to correct, our normalised similarity value verification method detected all the registration failures. Therefore no incorrect registrations would have been automatically displayed to the clinicians. Our discriminating criterion used would, however, have resulted in 14.49% of registrations requiring manual visual inspection. This is a large percentage of total registrations and so, although much better than the 100% which currently require manual visual inspection is still a long way from full automation. We propose that a combination of independent validation tests can greatly reduce manual input. We were unable to show the impact of combined tests on our 23 patient dataset (as specific data for VFM insertion was not collected). However we could show this impact on our evaluation dataset (8 patients) where the combined verification produced a TPR of 100% with a zero FPR.

We present in Figure 13 the standard pipeline of processes in a 2D-3D registration system alongside the pipeline when the novel methods proposed in this paper are used. Each processing step is represented by a separate box, and the box outlines denote the amount of manual input required for each step: a double outline with a thick outer border denotes knowledgeable manual input required; a double outline with a thin outer border denotes trivial manual input required; and a single thin border denotes no manual input required. The required knowledgeable input in the standard pipeline, i.e. *specific* vertebra identification and inspection of feature agreement in fluoro-DRR overlays, needs a high level of expertise, and is challenging even for an experienced operator. This is particularly true for fluoroscopy images picturing just a few vertebrae, where no landmark features such as the

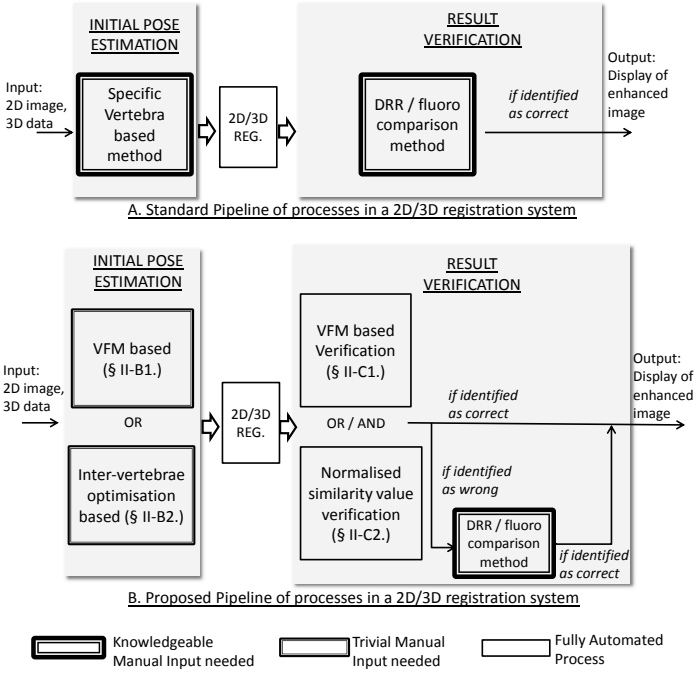

Fig. 13. The stages of a 2D/3D registration system when A. the standard methods and B. the proposed methods are used.

pelvis and the lowest rib are visible. On the other hand, the manual input required by our proposed methods for initial pose estimation is 1) VFM identification (equivalent to picking a point on a ruler) and fine tuning to the closest vertebra for the VFM based method and 2) detection of *any* vertebra on the fluoro for the inter-vertebrae optimisation based method. In both cases a much lower level of expertise is required. Moreover, no additional input is needed by the proposed methods for verification of registration. If the registration is not automatically verified then the standard method for manual verification (requiring knowledgable input) should be used, however, as discussed earlier, this will affect only a small minority of cases.

It is important to note that the initial pose estimation could be required many times during an operation (not just once at the beginning). For our system, any movements of the C-arm, require a re-registration. This requirement could be removed using machine-based 2D-3D registration [7]. However, this requires a high-end fluoroscopy system and so is not compatible with the majority of existing imaging equipment, especially mobile C-arms in standard operating theatres which are commonly used for endovascular procedures. In addition, even when using machine-based 2D-3D registration, patient movement can be a problem which requires re-registration. Our VFM based methods for initial position and result verification are also affected by patient movement. Significant patient movement would lead to increased errors in initial position estimation and a decrease in TPR. These could be corrected by repeating the process of VFM insertion, which, as it only requires two fluoroscopy images and a 20° gantry rotation, does not have large dose or time implications.

As discussed above, patient movement is a problem for machine-based 2D-3D registration. Before patient movement can be corrected it must first be detected. One possible use of our VFM verification method is to provide information on such patient movement. Figures 14 and 15 show how the VFM verification reprojection distance changes for each of the registrations carried out during the procedures in the evaluation data set. The patients are split into two groups, those who had a general anaesthetic (Figure 14) and those who had an epidural anaesthetic (Figure 15). The amount of movement can be seen to range from just a few mm up to 9mm. It is interesting how large amounts of patient movement are observed even during procedures where patients underwent a general anaesthetic. This is thought to be due to forces imparted by the clinicians when manipulating instruments over a long period of time. The large jumps in VFM reprojection distance observed in patients A and (to a lesser degree) C occurred when the surgeons changed from femoral to brachial artery access, which required movement of the patient's arm. Although machine-based 2D-3D registration is able to maintain high accuracy (less than 0.2mm [7]) this assumes zero patient movement. The movements observed by our VFM verification method would have seriously compromised the accuracy of a machine-based 2D-3D registration system, for the majority of procedures analysed.

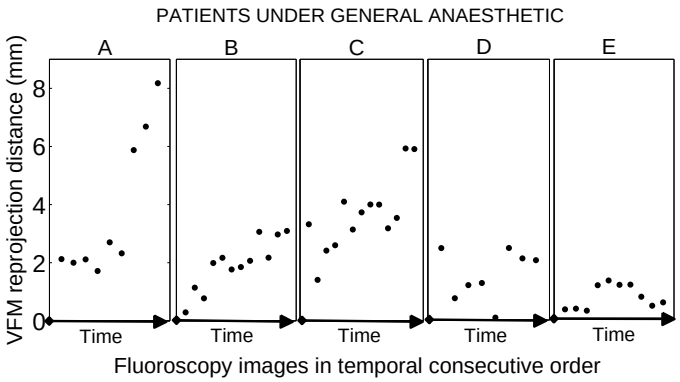

Fig. 14. VFM reprojection distance values shown over time shown per patient who underwent a general anaesthetic. Each dot represents an individual registration

Our proposed method to normalise the similarity measure value requires registration to a similar CT feature. Use of nearby CT vertebrae was natural for our vertebra-based system. A more generalised method to achieve similar CT features for any bony anatomy could be to use active appearance models [25] built from a population of example CT scans. Preoperatively the model could be registered to the patient's CT scan, but only allowing a very small number of modes of variation to be altered, so differences in shape and intensity between the patient's CT and model instantiation would remain. An alternative would be to warp the patient's CT slightly so the feature on which registration would be based would again have a slightly different shape to that observed in fluoroscopy. In both cases the similar CT feature should be able to normalise out the large variations in similarity measure caused by changes in fluoroscopy image quality and presence

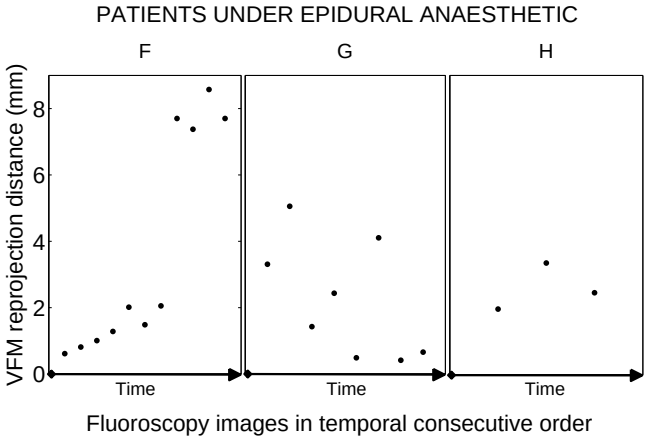

Fig. 15. VFM reprojection distance values shown over time shown per patient who had an epidural anaesthetic. Each dot represents an individual registration

of instruments.

As far as execution time is concerned, we are able to run a single vertebra registration in 13 seconds. This is 12 seconds for a serial CPU implementation of the production of the DRRs and 1 second for a massively parallel Graphics Processing Unit (GPU) implementation of the similarity measure of Eq. 5. The proposed normalised similarity value for verification requires the registration to take place for a number of candidate vertebrae. In clinical practice we use the method with 8 vertebrae (3 thoracic and 5 lumbar) on a system with a 4 cores CPU (Intel i7 960, 3.2GHz) and 2 GPUs (GeForce GTX 480). The total execution time is 28 seconds. Taking into account that a GPU implementation can render a DRR in 2msec [26], the production time of the DRRs in the registration algorithm can be reduced to at least 1.2 seconds (we consider approximately 600 DRRs per registration). Thus, an off-the-shelf system with 2 dual GPU Graphics Cards (4 GPUs in total) will perform the registration of 8 vertebrae in 4.4 seconds.

We believe that in the absence of a single "super" validation test method the combination of multiple independent high quality (but not perfect) test methods is a sensible method to achieve a good overally validation check. From our proposed methods the VFM provides information independent to the registration algorithm, while the normalised similarity measure depends on a major part of the registration algorithm : the similarity measure. This is why these two measures can be combined effectively to achieve a very high combined TPR. However, this means that a misclassification possibility will always exist, and so for clinical systems additional measures will be required. An alternative to presenting clinicians with only two classes (correct registrations or those requiring additional verification) would be to report a scalar confidence index on accuracy. Registrations resulting in low confidence could be verified using visual inspection, or an additional, slightly different registration (for example to a different vertebra, or after acquiring a higher quality "spot-film" fluoroscopy image) could be carried out until a required confidence value is achieved. Looking at the consistency between registrations and bringing imaging factors which are under clinician control into the validation process (e.g. widening field of view or providing "spot-film" images) could further reduce the requirement for visual inspection of results.

VI. Conclusions

We present methods to increase automation in initialisation and result verification for a 2D-3D image-guided surgery system. A retrospective study of the impact of the methods on the system, using data and starting positions acquired during 23 operations, showed the proposed methods significantly reduced the number of misregistrations and automatically verified the majority of results, rejecting all incorrect registrations.

Finally, let us point out again here that the potential importance of the proposed methods lies with their success in significantly decreasing the knowledgeable interaction between the clinicians and the registration system. This is a critical step towards the wider use of this technology in clinical practice. We strongly believe that the less input is required by the clinicians, the more such systems will be incorporated in clinical routine.

References

[1] P. Markelj, D. Tomaževič, B. Likar, and F. Pernuš, "A review of 3D/2D registration methods for image-guided interventions," *Med. Image Anal.*, vol. 16, no. 3, pp. 642–661, 2012.

[2] D. S. Fu and G. Kuduvalli, "A fast, accurate, and automatic 2D-3D image registration for image-guided cranial radiosurgery," *Med. Phys.*, vol. 35, no. 5, pp. 2180–2194, 2008.

[3] N. Agazaryan, S. E. Tenn, A. A. F. Desalles, and M. T. Selch, "Image-guided radiosurgery for spinal tumors: methods, accuracy and patient intrafraction motion," *Phys. Med. Biol.*, vol. 53, no. 6, pp. 1715–1727, 2008.

[4] A. A. Patel, P. Whang, and A. Vaccaro, "Overview of computer-assisted image-guided surgery of the spine," *Seminars in Spine Surgery*, vol. 20, no. 3, pp. 186–194, 2008.

[5] T. Tjardes, S. Shafizadeh, D. Rixen, T. Paffrath, B. Bouillon, E. S. Steinhausen, and H. Baethis, "Image-guided spine surgery: state of the art and future directions," *European Spine Journal*, vol. 19, pp. 25–45, 2010.

[6] K. Rhode and M. Sermesant, "Modeling and registration for electrophysiology procedures based on three-dimensional imaging," *Current Cardiovascular Imaging Reports*, vol. 4, pp. 116–126, 2011.

[7] D. Ruijters, R. Homan, P. Mielekamp, P. van de Haar, and D. Babic, "Validation of 3D multimodality roadmapping in interventional neuroradiology," *Phys. Med. Biol.*, vol. 56, pp. 5335–5354, 2011.

[8] D. Tomaževič, B. Likar, and F. Pernuš, "3-D/2-D registration by integrating 2-D information in 3-D," *IEEE Trans. Med. Imaging*, vol. 25, no. 1, pp. 17–27, 2006.

[9] P. Markelj, D. Tomaževič, F. Pernuš, and B. Likar, "Robust gradient-based 3-D/2-D registration of CT and MR to X-ray images," *IEEE Trans. Med. Imaging*, vol. 27, no. 12, pp. 1704–1714, 2008.

[10] D. Tomaževič, B. Likar, T. Slivnik, and F. Pernuš, "3-D/2-D registration of CT and MR to X-ray images," *IEEE Trans. Med. Imaging*, vol. 22, no. 11, pp. 1407–1416, 2003.

[11] E. B. van de Kraats, G. P. Penney, T. van Walsum, and W. J. Niessen, "Multispectral MR to X-ray registration of vertebral bodies by generating CT-like data," in *Medical Imaging Computing and Computer-Assisted Intervention—MICCAI '05*, 2005, pp. 911–918.

[12] X. Zhang, G. Y. Zheng, F. Langlotz, and L. P. Nolte, "Assessment of spline-based 2D-3D registration for image-guided spine surgery," *Minimally Invasive Therapy and Allied Technologies*, vol. 15, no. 3, pp. 193–199, 2006.

[13] S. Jonic, P. Thevenaz, and M. A. Unser, "Multispectral MR to x-ray registration of vertebral bodies by generating CT-like data," in *Medical Imaging 2003: Image Processing*, vol. 5032. San Diego, CA, USA: SPIE, 2003, pp. 1049–1052.

[14] D. B. Russakoff, T. Rohlfing, A. Ho, D. H. Kim, R. Shahidi, J. R. Adler, and C. R. Maurer, "Evaluation of intensity-based 2D-3D spine image registration using clinical gold-standard data," in *Second International Workshop on Biomedical Image Registration*, ser. Lecture Notes in Computer Science, vol. 2717. Philadelphia, PA, USA: Springer, 2003, pp. 151–160.
[15] H. Livyatan, Z. Yaniv, and L. Joskowicz, "Gradient-based 2-D/3-D rigid registration of fluoroscopic X-ray to CT," *IEEE Trans. Med. Imaging*, vol. 22, no. 11, pp. 1395–1406, 2003.
[16] J. Weese, G. P. Penney, P. Desmedt, T. M. Buzug, D. L. G. Hill, and D. J. Hawkes, "Voxel-based 2-D/3-D registration of fluoroscopy images and CT scans for image-guided surgery," *IEEE Trans. on Information Technology in Biomedicine*, vol. 1, no. 4, pp. 284–293, 1997.
[17] A. Hamadeh, S. Lavallee, and P. Cinquin, "Automated 3-dimensional computed tomographic and fluoroscopic image registration," *Comput. Aided Surg.*, vol. 3, no. 1, pp. 11–19, 1998.
[18] G. P. Penney, A. Varnavas, N. Dastur, and T. Carrell, "An image-guided surgery system to aid endovascular treatment of complex aortic aneurysms: Description and initial clinical experience," in *Information Processing in Computer-Assisted Interventions*, ser. Lecture Notes in Computer Science. Springer, 2011, pp. 13–24.
[19] G. P. Konin and D. M. Walz, "Lumbosacral transitional vertebrae: Classification, imaging findings, and clinical relevance," *American Journal of Neuroradiology*, vol. 33, no. 5, pp. 1778–1786, 2010.
[20] M. J. van der Bom, L. W. Bartels, M. A. Viergever, and J. P. W. Pluim, "Robust initialization of 2D–3D image registration using the projection-slice theorem and phase correlation," *Med. Phys.*, vol. 37, no. 4, pp. 1884–1892, 2010.
[21] C. Gendrin, P. Markelj, S. A. Pawiro, J. Spoerk, C. Bloch, C. Weber, M. Figl, H. Bergmann, W. Birkfellner, B. Likar, and F. Pernus, "Validation for 2D/3D registration. II: The comparison of intensity- and gradient-based merit functions using a new gold standard data set," *Med. Phys.*, vol. 38, no. 3, pp. 1491–1502, 2011.
[22] Y. Otake, M. Armand, R. Armiger, M. D. Kutzer, E. Basafa, P. Kazanzides, and R. H. Taylor, "Intraoperative image-based multiview 2D/3D registration for image-guided orthopaedic surgery: incorporation of fiducial-based C-arm tracking and GPU-acceleration," *IEEE Trans. Med. Imaging*, vol. 31, no. 4, pp. 948–962, 2012.
[23] G. P. Penney, J. Weese, P. G. Batchelor, D. L. G. Hill, and D. J. Hawkes, "Validation of a two- to three-dimensional registration algorithm for aligning preoperative CT images and intraoperative fluoroscopy images," *Med. Phys.*, vol. 28, no. 6, pp. 1024–1032, 2001.
[24] S. Theodoridis and K. Koutroumbas, Eds., *Pattern Recognition*, 4th ed. Academic Press, 2009, pp. 288–291.
[25] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 6, pp. 681–685, 2001.
[26] O. Dorgham, S. D. Laycock, and M. H. Fisher, "GPU accelerated generation of digitally reconstructed radiographs for 2D/3D image registration," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 9, pp. 2594–2603, 2012.

ACKNOWLEDGMENT

The authors would like to thank the Guy's and St Thomas' charity for funding this work. The authors also acknowledge financial support from the Department of Health via the National Institute for Health Research (NIHR) comprehensive Biomedical Research Centre award to Guy's & St Thomas' NHS Foundation Trust in partnership with King's College London and King's College Hospital NHS Foundation Trust. We would also like to thank all clinical staff at St Thomas' who aided this research and all of the patients who participated in this study.